3D visualization and analysis of HI in and around galaxies

Punzo, Davide

Summary

Astronomy is a visual science and heavily depends on data visualization. Data visualization provides a powerful tool both for interpreting the data and then for sharing the newly acquired knowledge with colleagues and the public. This concept was well known already during the time of the scientific revolution. In the 16th century, astronomers described the universe in their manuscripts with mathematical and visual representations. For example, Galileo Galilei described the universe as follows:

> *La filosofia è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi a gli occhi (io dico l'universo), ma non si può intendere se prima non s'impara a intender la lingua, e conoscer i caratteri, ne' quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi, ed altre figure geometriche, senza i quali mezi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro laberinto.*
>
> **– Galileo Galilei, Il Saggiatore (1623).**

i.e., translated in English, it states:

> *Philosophy is written in that great book which ever lies before our eyes — I mean the universe — but we cannot understand it if we do not first learn the language and grasp the symbols, in which it is written. This book is written in the mathematical language, and the symbols are triangles, circles and other geometrical figures, without whose help it is impossible to comprehend a single word of it; without which one wanders in vain through a dark labyrinth.*
>
> **– Galileo Galilei, The Assayer (1623), as translated by Thomas Salusbury in Mathematical collections and translations (1661).**

At that time, Galilei was used to illustrate his scientific discoveries with detailed drawings. As example, two of these, the observations of the moon phases (Sidereus Nuncius, 1610) and the orbits of the planets in the Copernican Solar system (Dialogue Concerning the Two Chief World Systems, 1632), are shown here:

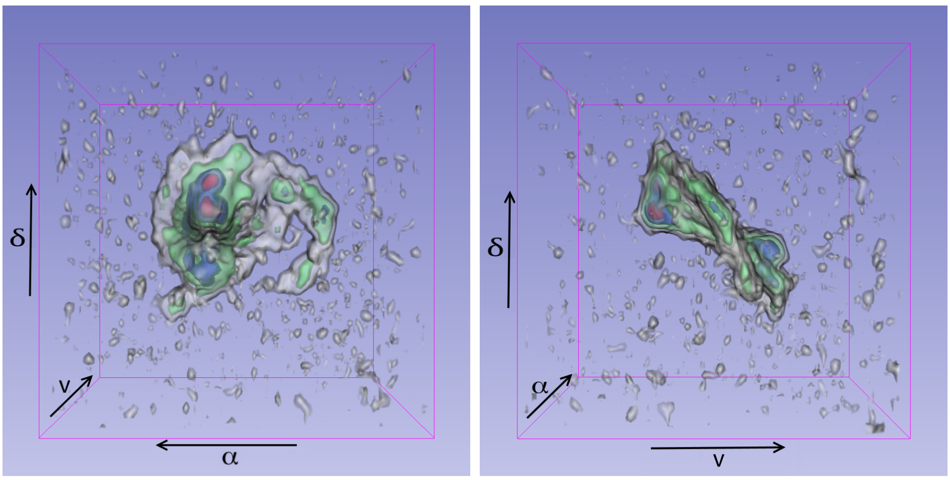
In the second half of the 20th century the advent of personal computers completely revolutionized the concept of data visualization and enabled its wide adoption, as it was made easier to access and use visualization tools by everyone interested. Data visualization was previously a specialized and time-consuming activity. This has been one of the factors that has stimulated the birth of new research fields such as scientific visualization and visual analytics. These fields focus on investigating algorithms to optimally visualize a wide range of types of data. Moreover the advent of *Graphics Processing Units* (GPUs), a technology supported and developed by the gaming industry, provided the computational power to achieve truly interactive visualization.

Interactivity is a key factor for enhancing data visualization in science, including astronomy, as interactive performance makes it possible to add analytics capabilities. Visual analysis (or Visual Analytics) software allows us not only to represent data graphically, but to also interact with those visual representations, to change the nature of the display, filter out what is not relevant, highlight lower levels of detail, and highlight subsets of data across multiple graphs simultaneously. This heavily involves the knowledge of the user and skills, such as pattern recognition and human assessment, in the analysis, resulting in insights that cannot be matched by traditional approaches. Static graphs delivered on paper or electronically on a computer screen help us communicate information in a clear and enlightening way, which is a benefit that should not be underestimated. It is, however, visual analytics solutions that help with the discovery of subtle components in complex and/or large datasets.

It is also worth remarking that data visualization is not about making data or plots look fancy. It is not about dressing up your presentations to dazzle your audience. Although delivering high quality images to show the beauty of celestial objects is a powerful way to engage interest in astronomy from the whole society, scientific visualization has little to do with art (though aesthetics are involved in displaying data effectively). Rather, it mostly involves science, a set of rules based on what we know about visual perception and cognition. When these rules are understood and software is designed to support them, then, and only then, it will positively affect and assist the researcher's daily work.

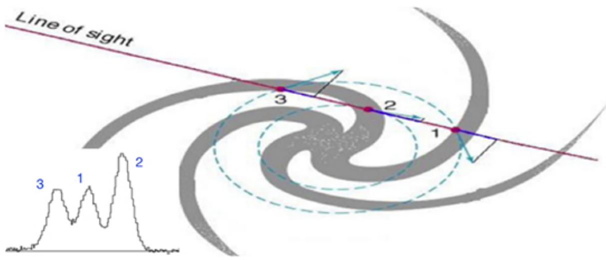
In my thesis, I reviewed the scientific requirements to visualize, inspect and analyze data of the neutral atomic hydrogen (H I) in galaxies that will be provided by large surveys planned with new or upgraded radio-telescopes, in particular, the Square Kilometre Array (SKA) precursors. One of them, Apertif, a new receiver system on the Westerbork Synthesis Radio Telescope (WSRT), is located in Netherlands and will provide a wealth of data regarding the gas content in hundreds of thousands of galaxies. Part of these data will provide H I signatures of gas falling into (and escaping from) galaxies. Studying these signatures is fundamental for a better understanding of the life cycle of galaxies. However, the H I emission is often very faint and its signature can be very subtle and complex to interpret. This makes it hard to efficiently identify them with automated software. Therefore, manual inspection and visualization of the sources will remain essential.

In the following example an observation of the galaxy WEIN069 in the Perseus-Pisces supercluster filament is used. The example shows in detail the character of the radio emission of H I in and around the galaxy:



The following link will redirect to a video of the rendering: `https://youtu.be/yLjW9nbdO8g`

In this image, a 3D view displays the H I distribution and the kinematics of WEIN069. The different colors highlight various intensity levels in the data. Grey, green, blue, and red correspond to 3, 8, 15, and 20 times the root mean square noise, respectively. The 3D view provides an immediate overview of the spatial and velocity structures in the data. The view confirms that the structures are coherent in all three dimensions. To interpret the 3D view one has to realize that it is a mix of spatial and velocity structure. Two spatial dimensions (right ascension, α, and declination, δ) measure the projection of H I onto the sky. The third radial velocity dimension measures the velocity of H I along the line of sight. The following schematic image shows a galaxy with velocity components 1, 2, and 3 along the line of sight and the profile of a schematic line along the line of sight:



The projection of the velocities along the line of sight accounts for the observation of these components at different velocities. For a given projection on the sky of a galaxy, geometric properties determine the precise projection of the velocity components, and hence the kinematic information of the observation. Such properties include the inclination and the position angle of the semi-major axis as well as the distribution of gas along the line of sight. The projection of the velocity components of a rotating system then determines the 3D shape.

In the volume rendering of the galaxy WEIN069, two main components are visible. The first component is a central part. This part represents the regularly rotating disk of the galaxy. Three-dimensional warping is caused by the projected rotation of the disk. The second component is a tail of unsettled gas (to the right of the main body in the first 3D view) that results from tidal interaction with another galaxy. The tail is connected to to the rotating disk in terms of space and velocity.

While 3D visualization facilitates the inspection of typical H I datasets, astronomers still require training to interpret the 3D structure of the data and to get acquainted with the tool itself. 3D visualization is a rather common tool in medical science, but misses a connection to astronomical data of the kind described in this thesis. Therefore, my research focused on the development of a new visualization tool which exploits interactive 3D visualization: `SlicerAstro`, a 3D visual analytic tool, based on traditional 2D hardware, for H I in and around galaxies. `SlicerAstro` is an open source extension of `3DSlicer`, a multi-platform open source software package for visualization and medical image processing. `SlicerAstro` does not only provide 3D visualization capabilities, but also a user-friendly Graphical User Interface (GUI) for controlling and interacting with the 3D displays. The following capabilities are available in the current distribution of `SlicerAstro`:

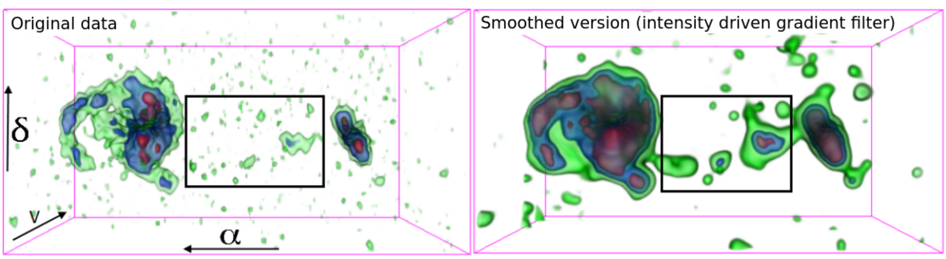
Interactive 3D visualization of astronomical data-cubes

`SlicerAstro` can import and save astronomical data in the Flexible Image Transport System (FITS) format. `SlicerAstro` can handle also the display of astronomical world coordinate systems (WCS). Once the data are loaded, `SlicerAstro` can fully exploit the `3DSlicer` interactive visualization capabilities.

Interactive 3D visualization enables an immediate overview of the coherent structures in the data, helping the inspection of complex datasets.

Interactive (semi)automatic smoothing in all three dimensions

`SlicerAstro` provides adaptive smoothing for H I data which runs on multi-core CPUs (`OpenMP`) and GPUs (`OpenGL`). This allows enhancing the local signal-to-noise ratio (S/N) of very faint structures ($S/N < 3$) such as tidal tails, ram-pressure tails, filaments and extraplanar gas. In the following example, I show that applying an adaptive filter (intensity driven gradient) to the data and then visualizing the new data-cube, the presence of a very faint filament between two galaxies is immediately spotted:

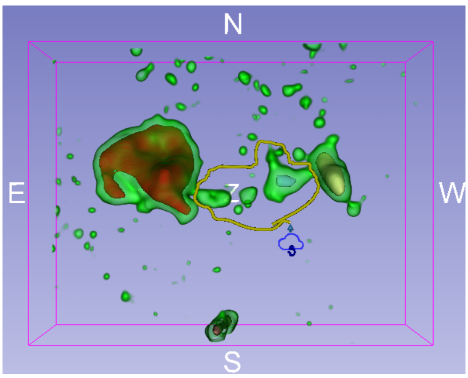


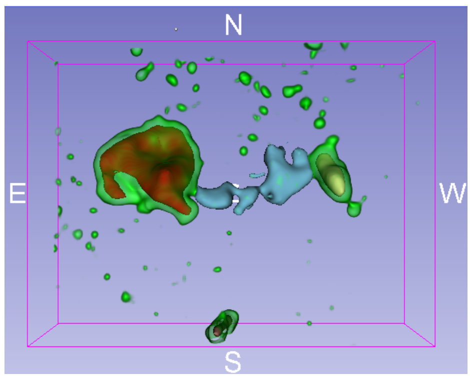
In these two images, the data are rendered in different colors that highlight the data at different intensity levels: green, blue and red correspond to 3, 7 and 15 times the *rms* noise respectively

Interactive semi-automatic filtering, such as the intensity driven gradient filter, coupled with interactive 3D visualization greatly helps the discovery of very faint signals.

Interactive 3D selection of H I sources

Volumetric data interaction tools (e.g., picking a voxel or selecting a region of interest in 3D) are necessary for performing data analysis in a 3D environment. The `AstroCloudLasso` selection tool available in `SlicerAstro` can perform selections in the 3D space by drawing directly in the 3D view as shown here:



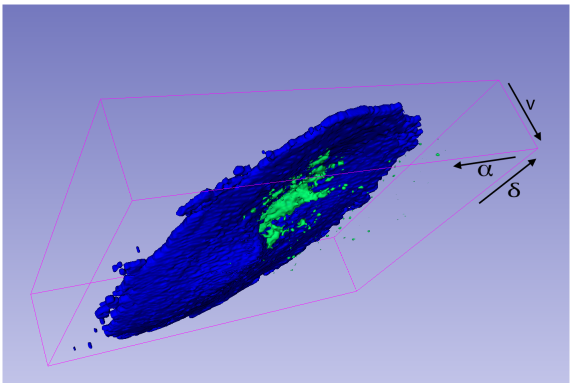


In these two images, the usage of the `AstroCloudLasso` segmentation editor effect is shown. A smoothed version of the WEIN069 data is rendered in green at the intensity level equal 3 times the *rms*. The colored segmentations (red, yellow, brown and cyan) represent the masks generated by an automated algorithm. These can be modified (or new ones can be created) by the `AstroCloudLasso` selection tool. The selection is visualized as a yellow tube (i.e, 2D lasso) drawn by the user with the 2D cursor indicated by the blue cloud. This tool computes a selection in 3D space from the 2D user-selection. It builds a closed surface at the value of the intensity level specified (in the GUI) and then visualizes the modified segment as shown in the second image.

This selection tool can be used for: i) selecting regions of interest for further analysis; ii) modifying masks (e.g., objects created by automated algorithms which identify the extent of the sources in the data-cube).

Interactive H I data modeling coupled to visualization

Interactively visualizing kinematic models (e.g., tilted-ring models that describe the rotation and geometry of H I) and data can help highlight asymmetries, extra-planar gas, and other special features. Interactively changing and displaying the model parameters, for instance, can provide more effective comparisons of models and data. One powerful possibility is to use the output model of an automated model-fitting algorithm to visually highlight different components in a data cube. For example, the image below displays data from the galaxy NGC2403 (in green) and a model that fits the disk of the galaxy (in blue):



The image clearly shows the presence of an extra gas component (green) not fitted by the model (blue). This component consists of extraplanar gas close to the disk of the galaxy that rotates at a lower velocity than the gas in the disk.

Thus, the combined 3D visualization of data and models gives an immediate overview of different structures in the galaxy.

The next development for `SlicerAstro` is the addition of the Virtual Observatory (VO) communication protocols. This infrastructure allows to connect astronomical software packages and to easily share data (and results from the analysis) to colleagues. This connection will be very important to enhance the discovery of the unexpected in the era of Big Data of radio astronomy.

Finally, although the development of `SlicerAstro` thus far mainly focused on 3D H I data, it will also be a useful tool for any other type of 3D astronomical data such as mm/submm molecular line data and optical integral field spectroscopic data.